

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Oon, Chong Jin
Lim, Gek Keow
Leong, Ai Lin
Zhao, Yi
Chen, Wei Ning

(ii) TITLE OF INVENTION: A VACCINE-INDUCED HEPATITIS B VIRAL
STRAIN AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 11

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Ladas & Parry
(B) STREET: 26 West 61 Street
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: USA
(F) ZIP: 10023

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/719,533
(B) FILING DATE: 13-DEC-2000
(C) CLASSIFICATION: 435

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/SG98/00045
(B) FILING DATE: 19-JUN-1998

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Mass, Clifford J.
(B) REGISTRATION NUMBER: 30,086
(C) REFERENCE/DOCKET NUMBER: U-013108-9

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (212) 708-1890

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3215 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCCACCACT	TTCCACCAAA	CTCTTCAAGA	TCCCAGAGTC	AGGGCCCTGT	ACTTTCCTGC	50
TGGTGGCTCC	AGTTCAGGAA	CAGTGAGCCC	TGCTCAGAAT	ACTGTCTCTG	CCATATCGTC	100
AATCTTATCG	AAGACTGGGG	ACCCTGTACC	GAACATGGAG	AACATCGCAT	CAGGACTCCT	150
AGSACCCCTG	CTCGTGTTAC	AGGCGGGGTT	TTTCTTGTTG	ACAAAAATCC	TCACAATAAC	200
GCAGAGTCTA	GACTCGTGGT	GGACTTCTCT	CAATTTTCTA	GGGGGAACAC	CCGTGTGTCT	250
TGGGCAAAAT	TGSCAGTCCC	AAATCTCCAG	TCACTCACCA	ACCTGTTGTC	CTCCAAATTG	300
TGCTGGTTAT	CGGTGGATGT	GTGTGCGGCG	TTTTATCATC	TTCCTCTGCA	TGCTGTGTCT	350
ATGCTTCATC	TTCTTGTGGG	TTCTTCTGGA	CTATCAAGGT	ATGTTGCCCC	TTTGTCTCTT	400
AATTCCAGGA	TCAACAACAA	CCAGCACCGG	ACCATGCAAA	ACCTGCACAA	CTCCTGTCTA	450
AGGAACCTCT	ATGTTTCCCT	CATGTTGCTG	TACAAAACCT	ACGGACAGAA	ACTGCACCTG	500
TATTCCCATC	CCATCATCTT	GGGCTTTGCG	AAAATACCTA	TGGGAGTGGG	CCTCAGTCCG	550
TTTCTCTTGG	CTCAGTTTAC	TAGTGCCATT	TGTTCACTGG	TTCGTAGGGC	TTTCCCCCAC	600
TGTCTGGCTT	TCAGTTATAT	GGATGATGTG	GTTTTGGGGG	CCAAGTCTGT	ACAACATCTT	650
CGTCCCTTTT	ATGCCGCTGT	TACCAATTTT	CTTTTGTCTT	TGGGTATACA	TTTAAACCTT	700
CACAAAACAA	AAAGATGGGG	ATATTCCCTT	AACTTCATGG	GATATGTCAT	TGGGAGTTGG	750
GGCACATTGC	CACAGGAACA	TATTGTACAA	AAAATCAAAA	TGTGTTTTAG	GAAACTTCCT	800
GTAAACAGGC	CTATTGATTG	GAAAGTATGT	CAACGAATTG	TGGGTCTTTT	GGGGTTTGCC	850
GCCCCCTTCA	CGCAATGTGG	ATATCCTGCT	TTAATGCCTT	TATATGCATG	TATACAAGCA	900
AAACAGGCTT	TTACTTTCTC	GCAAACTTAC	AAGACCTTTC	TAAGTAAACA	GATCTGTAAC	950
CTTTACCCCG	TTGCTGGGCA	ACGCCCTGGT	CTGTGCCAAG	TGTTTGCTGA	CGCAACCCCG	1000
ACTGGTTGGG	GCTTGGCCAT	AGGCCATCAG	CGCATGCGTG	GAACCTTTGT	GTCTCTCTGT	1050
CGGATCCATA	CTGGGGAAC	CCTAGCCGCT	TGTTTTGCTC	GCAGCAGGTC	TGGGGCAAAA	1100
CTCATCGGGA	CTGACAATTG	TGTCGTGCTC	TCCCACAAGT	ATACATCATT	TGCATGGGTG	1150
CTAGGCTGTG	CTGCCAACTG	GATCTTGCGC	GGGACGTCCT	TTGTTTACGT	CCCGTGGGGG	1200
CTGAATCCCG	CGGACGACCC	CTCCCGGGGC	CGCTTGGGGC	TCTACCGCCC	GCTTCTCCCG	1250
CTGTTATACC	GACCGACCAC	GGGGCGCACC	TCTCTTTACG	CGGACTCCCC	GTCTGTGCTT	1300
TCTCATCTGC	CGGACCGTGT	GCACTTCGCT	TCACCTCTGC	ACGTCGCATG	GAGACCAACG	1350
TGAACGCCCC	CGGGAACCTG	CCCAAGGTCT	TGCATAAGAG	GACTCTTGGA	CTTTTCAGCAA	1400

TGTCAACGAC	CGACCTTGAG	GCATACTTCA	AAGACTGTGT	GTTTAATGAG	TGGGAGGAGT	1740
TGGGGGAGGA	GGTTAGGTTA	AAGGTCTTTG	TACTAGGAGG	CTGTAGGCAT	AAATTGGTGT	1800
GTTCAACATC	AACATGCAAC	TTTTTCACCT	CTGCCTAATC	ATCTCATGTT	CATGTCTTAC	1860
TGTTCAAGCC	TCCAAGCTGT	GCCTTGGGTG	GCTTTGGGGC	ATGGACATTG	ACCCGTATAA	1920
AGAATTTGGA	GCTTCTGTGG	AGTTACTCTC	TTTTTTGCCT	TCTGACTTTT	TTCTTCTAT	1980
TGAGATCTC	CTCGACACCG	CCTCTGCTCT	GATCGGGAG	GCCTTAGAGT	CTCCGGAACA	2040
TTGTTACCT	CACCATACGG	CACTCAGGCA	AGCTATTCTG	AGTTGGGGTG	AGTTAATGAA	2100
TCTAGCCACC	TGGGTGGGAA	GTAATTTGGA	AGATCCAGCA	TCCAGGGAAT	TAGTAGTCAG	2160
CTATGTCAAC	GTTAATATGG	GCCTAAAAAT	CAGACAACTA	TTGTGGTTTC	ACATTTCTTG	2220
TCTTACTTTT	GGGAGAGAAA	CTGTTCTTGA	ATATTTGGTG	TCTTTTGGAG	TGTGGATTGG	2280
CACTCCTCCT	GCATATAGAC	CACCAAATGC	CCCTATCTTA	TCAACACTTC	CGGAAACTAC	2340
TGTTGTTAGA	CGAAGAGGCA	GGTCCCCTAG	AAGAAGAACT	CCCTCGCCTC	GCAGACGAAG	2400
GTCTCAATCG	CGCGCTCGCA	GAAGATCTCA	ATCTCGGGAA	TCTCAATGTT	AGTATTCCTT	2460
GGACACATAA	GGTGGGAAAC	TTTACGGGGC	TTTATTCTTC	TACGSTACCT	TGCTTTAATC	2520
CTAAATGGCA	AACTCCTTCT	TTTCCGGACA	TTCATTTGCA	GGAGGACATT	CTTGATAGAT	2580
GTAAGCAATT	TGTGGGGCCC	CTTACAGTAA	ATGAAAACAG	GAGACTAAAA	TTAATTATGC	2640
CTCTAGGTT	TTATCCAAAT	GTTACTAAAT	ATTTGCCCTT	AGATAAAGGG	ATCAAACCAAT	2700
AATATCCAGA	GATGTAGTT	AATCATTACT	TCCAGACGCG	ACATTATTTA	CACACTCTTT	2760
CGAAGGCGGG	GATCTTATAT	AAAAGAGAGT	CCACACGTAG	CGCCTCATTT	TGCGGGTCACT	2820
CATAATCTTG	GGAAACAAGAT	CTACAGCATG	GGAGGTTGGT	CTTCCAAACC	TGAAAAGGC	2880
ATGGGACAA	ATCTTTCTGT	CTCCAATCCC	CTGGGATTCT	TCCCAGATCA	TGAGTTGGAT	2940
CTTGCAATCA	AAACCAACTC	AGAAAATCCA	GATTGGGACC	TCAACCCGCA	CAAGGACAAC	3000
TGGCGGGACG	CCAACAAGGT	GGGAGTGGGA	GCATTCGGGC	CAGGTTTCAC	CCCTCTCTAT	3060
GGGGACTGT	TGGGTGGAG	CCCTCAGGCT	CAGGGCCTAC	TCACAACGTG	GCCAGCAGCT	3120
CTCTCTCTTG	CTTCAACCAA	TGGCAGTCA	GGAAGGCAGC	CTACTCCCTT	ATCTCCAGCT	3180
CTAAAGGACA	CTCATCTTCA	GGCATGCA	TGGAA			3215

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 843 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Pro	Leu	Ser	Tyr	Gln	His	Phe	Arg	Lys	Leu	Leu	Leu	Leu	Asp	Glu	
1			5						10					15		
Glu	Ala	Gly	Pro	Leu	Glu	Glu	Glu	Leu	Pro	Arg	Leu	Ala	Asp	Glu	Gly	
			20					25					30			
Leu	Asn	Arg	Arg	Val	Ala	Glu	Asp	Leu	Asn	Leu	Gly	Asn	Leu	Asn	Val	
		35					40					45				
Ser	Ile	Pro	Trp	Thr	His	Lys	Val	Gly	Asn	Phe	Thr	Gly	Leu	Tyr	Ser	
	50					55					60					
Ser	Thr	Val	Pro	Cys	Phe	Asn	Pro	Lys	Trp	Gln	Thr	Pro	Ser	Phe	Pro	
65					70					75					80	
Asp	Ile	His	Leu	Gln	Glu	Asp	Ile	Leu	Asp	Arg	Cys	Lys	Gln	Phe	Val	
			85						90					95		
Glu	Pro	Leu	Thr	Val	Asn	Glu	Asn	Arg	Arg	Leu	Lys	Leu	Ile	Met	Pro	
			100					105					110			
Ala	Arg	Phe	Tyr	Pro	Asn	Val	Thr	Lys	Tyr	Leu	Pro	Leu	Asp	Lys	Gly	
		115					120					125				
Ile	Lys	Pro	Tyr	Tyr	Pro	Glu	Tyr	Val	Val	Asn	His	Tyr	Phe	Gln	Thr	
	130					135					140					
Arg	His	Tyr	Leu	His	Thr	Leu	Trp	Lys	Ala	Gly	Ile	Leu	Tyr	Lys	Arg	
145					150					155					160	
Glu	Ser	Thr	Arg	Ser	Ala	Ser	Phe	Cys	Gly	Ser	Pro	Tyr	Ser	Trp	Glu	
				165					170					175		
Gln	Asp	Leu	Gln	His	Gly	Arg	Leu	Val	Phe	Gln	Thr	Ser	Lys	Arg	His	
		180						185					190			
Gly	Asp	Lys	Ser	Phe	Cys	Pro	Glu	Ser	Pro	Gly	Ile	Leu	Pro	Arg	Ser	
		195					200					205				
Ser	Val	Gly	Pro	Cys	Ile	Gln	Ser	Gln	Leu	Arg	Lys	Ser	Arg	Leu	Gly	
	210					215					220					
Pro	Gln	Pro	Ala	Gln	Gly	Gln	Leu	Ala	Gly	Arg	Gln	Gln	Gly	Gly	Ser	
225				230						235					240	

Gly Ser Ile Arg Ala Arg Val His Pro Ser Ser Trp Gly Thr Val Gly
 245 250 255
 Val Glu Pro Ser Gly Ser Gly Pro Thr His Asn Cys Ala Ser Ser Ser
 260 265 270
 Ser Ser Cys Leu His Gln Ser Ala Val Arg Lys Ala Ala Tyr Ser Leu
 275 280 285
 Ile Ser Thr Ser Lys Gly His Ser Ser Ser Gly His Ala Val Glu Leu
 290 295 300
 His His Phe Pro Pro Asn Ser Ser Arg Ser Gln Ser Gln Gly Pro Val
 305 310 315 320
 Leu Ser Cys Trp Trp Leu Gln Phe Arg Asn Ser Glu Pro Cys Ser Glu
 325 330 335
 Tyr Cys Leu Cys His Ile Val Asn Leu Ile Glu Asp Trp Gly Pro Cys
 340 345 350
 Thr Glu His Gly Glu His Arg Ile Arg Thr Pro Arg Thr Pro Ala Arg
 355 360 365
 Val Thr Gly Gly Val Phe Leu Val Asp Lys Asn Pro His Asn Thr Ala
 370 375 380
 Glu Ser Arg Leu Val Val Asp Phe Ser Gln Phe Ser Arg Gly Asn Thr
 385 390 395 400
 Arg Val Ser Trp Pro Lys Phe Ala Val Pro Asn Leu Gln Ser Leu Thr
 405 410 415
 Asn Leu Leu Ser Ser Asn Leu Ser Trp Leu Ser Leu Asp Val Ser Ala
 420 425 430
 Ala Phe Tyr His Leu Pro Leu His Pro Ala Ala Met Pro His Leu Leu
 435 440 445
 Val Gly Ser Ser Gly Leu Ser Arg Tyr Val Ala Arg Leu Ser Ser Asn
 450 455 460
 Ser Arg Ile Asn Asn Asn Glu His Arg Thr Met Glu Asn Leu His Asn
 465 470 475 480
 Ser Cys Ser Arg Asn Leu Tyr Val Ser Leu Met Leu Leu Tyr Lys Thr
 485 490 495
 Tyr Gly Gln Lys Leu His Leu Tyr Ser His Pro Ile Ile Leu Gly Phe
 500 505 510
 Arg Lys Ile Pro Met Gly Val Gly Leu Ser Pro Phe Leu Leu Ala Gln
 515 520 525
 Phe Thr Ser Ala Ile Cys Ser Val Val Arg Arg Ala Phe Pro His Cys

530	535	540
Leu Ala Phe Ser Tyr Met Asp Asp Val Val Leu Gly Ala Lys Ser Val		
545	550	555 560
Gln His Leu Glu Ser Leu Tyr Ala Ala Val Thr Asn Phe Leu Leu Ser		
	565	570 575
Leu Gly Ile His Leu Asn Pro His Lys Thr Lys Arg Trp Gly Tyr Ser		
	580	585 590
Leu Asn Phe Met Gly Tyr Val Ile Gly Ser Trp Gly Thr Leu Pro Gln		
	595	600 605
Glu His Ile Val Gln Lys Ile Lys Met Cys Phe Arg Lys Leu Pro Val		
	610	615 620
Asn Arg Pro Ile Asp Trp Lys Val Cys Gln Arg Ile Val Gly Leu Leu		
	625	630 635 640
Gly Phe Ala Ala Pro Phe Thr Gln Cys Gly Tyr Pro Ala Leu Met Pro		
	645	650 655
Leu Tyr Ala Cys Ile Gln Ala Lys Gln Ala Phe Thr Phe Ser Gln Thr		
	660	665 670
Tyr Lys Thr Phe Leu Ser Lys Gln Tyr Leu Asn Leu Tyr Pro Val Ala		
	675	680 685
Arg Gln Arg Pro Gly Leu Cys Glu Val Phe Ala Asp Ala Thr Pro Thr		
	690	695 700
Gly Trp Gly Leu Ala Ile Gly His Gln Arg Met Arg Gly Thr Phe Val		
	705	710 715 720
Ser Pro Leu Pro Ile His Thr Ala Glu Leu Leu Ala Ala Cys Phe Ala		
	725	730 735
Arg Ser Arg Ser Glu Ala Lys Leu Ile Gly Thr Asp Asn Ser Val Val		
	740	745 750
Leu Ser Arg Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly Cys Ala Ala		
	755	760 765
Asn Trp Ile Leu Arg Gly Thr Ser Phe Val Tyr Val Pro Ser Ala Leu		
	770	775 780
Asn Pro Ala Asp Asp Pro Ser Arg Gly Arg Leu Gly Leu Tyr Arg Pro		
	785	790 795 800
Leu Leu Arg Leu Leu Tyr Arg Pro Thr Thr Gly Arg Thr Ser Leu Tyr		
	805	810 815
Ala Asp Ser Pro Ser Val Pro Ser His Leu Pro Asp Arg Val His Phe		

820

825

830

Ala Ser Pro Leu His Val Ala Trp Arg Pro Pro
 835 840

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Gly	Gly	Trp	Ser	Ser	Lys	Pro	Arg	Lys	Gly	Met	Gly	Thr	Asn	Leu	1	5	10	15
Ser	Val	Pro	Asn	Pro	Leu	Gly	Phe	Phe	Pro	Asp	His	Gln	Leu	Asp	Pro	20	25	30	
Ala	Phe	Lys	Ala	Asn	Ser	Glu	Asn	Pro	Asp	Trp	Asp	Leu	Asn	Pro	His	35	40	45	
Lys	Asp	Asn	Trp	Pro	Asp	Ala	Asn	Lys	Val	Gly	Val	Gly	Ala	Phe	Gly	50	55	60	
Pro	Gly	Phe	Thr	Pro	Pro	His	Gly	Gly	Leu	Leu	Gly	Trp	Ser	Pro	Gln	65	70	75	80
Ala	Gln	Gly	Leu	Leu	Thr	Thr	Val	Pro	Ala	Ala	Pro	Pro	Pro	Ala	Ser	85	90	95	
Thr	Asn	Arg	Gln	Ser	Gly	Arg	Gln	Pro	Thr	Pro	Leu	Ser	Pro	Pro	Leu	100	105	110	
Arg	Asp	Thr	His	Pro	Gln	Ala	Met	Gln	Trp	Asn	Ser	Thr	Thr	Phe	His	115	120	125	
Gln	Thr	Leu	Gln	Asp	Pro	Arg	Val	Arg	Ala	Leu	Tyr	Phe	Pro	Ala	Gly	130	135	140	
Gly	Ser	Ser	Ser	Gly	Thr	Val	Ser	Pro	Ala	Gln	Asn	Thr	Val	Ser	Ala	145	150	155	160
Ile	Ser	Ser	Ile	Leu	Ser	Lys	Thr	Gly	Asp	Pro	Val	Pro	Asn	Met	Glu	165	170	175	
Asn	Ile	Ala	Ser	Gly	Leu	Leu	Gly	Pro	Leu	Leu	Val	Leu	Gln	Ala	Gly	180	185	190	

Phe Phe Leu Leu Thr Lys Ile Leu Thr Ile Pro Gln Ser Leu Asp Ser
 195 200 205
 Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Pro Thr Val Cys Leu Gly
 210 215 220
 Gln Asn Ser Gln Ser Gln Ile Ser Ser His Ser Pro Thr Cys Cys Pro
 225 230 235 240
 Pro Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile
 245 250 255
 Phe Leu Cys Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu
 260 265 270
 Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Thr
 275 280 285
 Thr Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr Thr Pro Ala Gln Gly
 290 295 300
 Thr Ser Met Phe Pro Ser Cys Cys Cys Thr Lys Pro Thr Asp Arg Asn
 305 310 315 320
 Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Ala Lys Tyr Leu
 325 330 335
 Trp Glu Trp Ala Ser Val Arg Phe Ser Trp Leu Ser Leu Leu Val Pro
 340 345 350
 Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu Ser Val
 355 360 365
 Ile Trp Met Met Trp Phe Trp Gly Pro Ser Leu Tyr Asn Ile Leu Ser
 370 375 380
 Pro Phe Met Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr Ile
 385 390 395 400

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 212 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gln Leu Phe His Leu Cys Leu Ile Ile Ser Cys Ser Cys Pro Thr
 1 5 10 15

Val	Gln	Ala	Ser	Lys	Leu	Cys	Leu	Gly	Trp	Leu	Trp	Gly	Met	Asp	Ile		
			20					25					30				
Asp	Pro	Tyr	Lys	Glu	Phe	Gly	Ala	Ser	Val	Glu	Leu	Leu	Ser	Phe	Leu		
		35					40					45					
Pro	Ser	Asp	Phe	Phe	Pro	Ser	Ile	Arg	Asp	Leu	Leu	Asp	Thr	Ala	Ser		
		50				55					60						
Ala	Leu	Tyr	Arg	Glu	Ala	Leu	Glu	Ser	Pro	Glu	His	Cys	Ser	Pro	His		
65					70					75					80		
His	Thr	Ala	Leu	Arg	Gln	Ala	Ile	Leu	Ser	Trp	Gly	Glu	Leu	Met	Asn		
			85					90						95			
Leu	Ala	Thr	Trp	Val	Gly	Ser	Asn	Leu	Glu	Asp	Pro	Ala	Ser	Arg	Glu		
			100					105						110			
Leu	Val	Val	Ser	Tyr	Val	Asn	Val	Asn	Met	Gly	Leu	Lys	Ile	Arg	Gln		
		115					120					125					
Leu	Leu	Trp	Phe	His	Ile	Ser	Cys	Leu	Thr	Phe	Gly	Arg	Glu	Thr	Val		
		130				135					140						
Leu	Glu	Tyr	Leu	Val	Ser	Phe	Gly	Val	Trp	Ile	Arg	Thr	Pro	Pro	Ala		
145					150					155					160		
Tyr	Arg	Pro	Pro	Asn	Ala	Pro	Ile	Leu	Ser	Thr	Leu	Pro	Glu	Thr	Thr		
			165					170						175			
Val	Val	Arg	Arg	Arg	Gly	Arg	Ser	Pro	Arg	Arg	Arg	Thr	Pro	Ser	Pro		
			180					185					190				
Arg	Arg	Arg	Arg	Ser	Gln	Ser	Pro	Arg	Arg	Arg	Arg	Ser	Gln	Ser	Arg		
		195					200					205					
Glu	Ser	Gln	Cys														
		210															

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ala	Ala	Arg	Leu	Cys	Cys	Gln	Leu	Asp	Pro	Ala	Arg	Asp	Val	Leu		
1				5					10					15			
Cys	Leu	Arg	Pro	Val	Gly	Ala	Glu	Ser	Arg	Gly	Arg	Pro	Leu	Pro	Gly		
			20					25					30				

Pro	Leu	Gly	Ala	Leu	Pro	Pro	Ala	Ser	Pro	Pro	Val	Ile	Pro	Thr	Asp
	35						40					45			
His	Gly	Ala	His	Leu	Ser	Leu	Arg	Gly	Leu	Pro	Val	Cys	Ala	Phe	Ser
	50					55					60				
Ser	Ala	Gly	Pro	Cys	Ala	Leu	Arg	Phe	Thr	Ser	Ala	Arg	Arg	Met	Glu
	65				70					75				80	
Thr	Thr	Val	Asn	Ala	His	Gly	Asn	Leu	Pro	Lys	Val	Leu	His	Lys	Arg
			85						90					95	
Thr	Leu	Gly	Leu	Ser	Ala	Met	Ser	Thr	Thr	Asp	Leu	Glu	Ala	Tyr	Phe
			100					105					110		
Lys	Asp	Cys	Val	Phe	Asn	Glu	Trp	Glu	Glu	Leu	Gly	Glu	Glu	Val	Arg
		115					120					125			
Leu	Lys	Val	Phe	Val	Leu	Gly	Gly	Cys	Arg	His	Lys	Leu	Val	Cys	Ser
	130					135					140				
Pro	Ser	Pro	Cys	Asn	Phe	Phe	Thr	Ser	Ala						
	145				150										

(2) INFORMATION FOR SEQ ID NO:6:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATAAGCTTATG CCGCTATCTT ATCAACACTT CCGGA

35

(2) INFORMATION FOR SEQ ID NO:7:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGTCTAGAC TCTGCGGTAT TGTGA

25

(2) INFORMATION FOR SEQ ID NO:8:

- (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAGTCTAGAC TCGTGGTGA CTTCT

25

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGAGAATTCT CACGGTGGTC TCCATGCGAC GT

32

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTTGTTTACG TCCCGT

16

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTAAGCTTAG TTTCCGGAAG TGTTGAT